Published in final edited form as:

Curr Opin Chem Biol. 2009 June ; 13(3): 345–353. doi:10.1016/j.cbpa.2009.04.637.

G-quadruplex nucleic acids as therapeutic targets

Shankar Balasubramanian^{1,3} and Stephen Neidle²

¹Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

²Cancer Research UK Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX, UK

³School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SP, UK

Abstract

Nucleic acid sequences containing several short runs of guanine nucleotides can form complex higher order structures, termed quadruplexes. Their occurrence has been most extensively characterised at the telomeric ends of eukaryotic chromosomes, whose DNA comprises such sequences, and where the extreme 3' ends are single-stranded. This enables relatively facile formation of quadruplex arrangements under the influence of a quadruplex-selective small molecule to compete effectively with telomeric protein-DNA interactions. Occurrences of quadruplexes within the human and other genomes have been mapped by bioinformatics surveys, which have revealed over-representations in promoter regions, especially of genes involved in replication, such as oncogenes, as well as in 5' UTR regions. The highly distinctive nature of quadruplex topologies suggests that they can act as novel therapeutic targets, for example in the selective inhibition of transcription of a given oncogene, using designed small molecules to stabilise a particular quadruplex. This offers the prospect of an alternative to, for example, direct kinase targeting with small molecules, without the attendant issues of active-site resistance. We survey here the basis of these approaches, together with current progress, and discuss the mechanistic issues posed by quadruplex targeting.

Introduction

The majority of cytotoxic cancer chemotherapeutic agents target DNA, in a relatively indiscriminate manner. The effectiveness of agents such as *cis*-platinum and Adriamycin is a consequence of DNA-repair defects and elevated DNA topoisomerase II levels, respectively, in susceptible cancer cell types. However these features, though highly significant for the positive clinical outcomes sometimes seen with these agents, are counter-balanced by their high toxicity and generation of resistance mechanisms. There has been little development of new cytotoxic drugs over the past few years, contrasting remarkably with the major effort world-wide in the discovery and the development of targeted agents that can exploit the knowledge of the molecular basis of cancer. However the latter have most effect against cancers that have a relatively simple biology, exemplified by the success of Gleevec® against the tyrosine kinase domain of the abl kinase in chronic myelogenous leukaemia (CML), where the bcr-abl chromosomal translocation is a determinant of this disease [1]. Kinase inhibitors and other targeted molecules show much less effect in the more common solid tumours; even in CML, resistance to Gleevec® is of major significance as a result of induced mutations in the ATP binding site of the abl kinase. The paradox is that cytotoxic agents continue to be of major importance in the current management of these cancers, even though their therapeutic window remains very narrow [2].

So how can DNA be a more selective target in cancer? There has been much effort towards the design of DNA sequence-specific molecules that can modulate transcription of particular genes; even if these targets are still kinases, such molecules would not have the disadvantage of generating resistance at the kinase level at least. However sequence-specific molecules have to recognise sequences of ca 16-20 nucleotides, and such molecules, though now available through the ingenuity of the polyamide approach pioneered by Dervan and co-workers [3], have not yet emerged as potential therapeutic agents, due to various reasons. We describe here an alternative approach to target specific DNA sequences that may combine high target selectivity with the prospect of developing small-drug-like molecules. This entails the targeting of DNA sequence motifs that fold into four-stranded structures called G-quadruplexes, with characteristics that are particular to the sequence involved. An early example of a G-quadruplex-targeting small-molecule drug has recently entered phase II clinical trials (see www.cylenepharma.com) and several others are likely to follow.

General features of quadruplex nucleic acids

G-quadruplex nucleic acids arise from tandem repeats of short guanine tracts; the guanines form G-quartet motifs that are held together by the intervening sequences (loops) [4-6]. A general G-quadruplex-forming sequence can be defined as:

where N_{L1-3} are loops that vary in length and sequence. The G-tracts form the core of a quadruplex, as a set of stacked G-quartets, with the loops positioned on the exterior, helping to hold the overall structure intact. The spaces between the loops, termed grooves, are bounded by charged phosphodiester backbones. Intramolecular G-quadruplexes can fold in a variety of ways, and the loops themselves can have elements of secondary structure even when comprising as few as three nucleotides. Loops can be diagonal, lateral or chain-reversal (also termed propeller), and the presence of a particular loop type in a structure is dependent on the number of G-quartets comprising a quadruplex, on loop length and on sequence. It was assumed until recently that the loops and core have separate roles in G-quadruplex architecture, but the NMR structure determination [7] of a G-quadruplex from the *c-kit* promoter (see below) showed that a guanine residue within a loop can also be an integral part of the G-quartet core. Such added complexities make G-quadruplex structure prediction challenging for all but the simplest examples.

All G-quadruplexes have four distinct phosphodiester chains, which can be mutually parallel or anti-parallel, dependent on the type of loop connecting them. Few of the possible topologies have been observed to date by NMR or crystallography, either as native structures or as small-molecule complexes, although some therapeutically relevant starting-points for structure-based drug design are now available, for human telomeric [8-11,12•,13], *c-myc* [14], *c-kit* [7] and *bcl-2* [15] quadruplexes. All quadruplex structures are very distinctive from duplex nucleic acids, offering considerable potential for differential molecular recognition, and thus have enabled a number of small molecules to be developed that have much higher quadruplex compared to duplex affinity [16,17]. This is essential if such molecules are to be used as biological G-quadruplex-directed probes or considered for therapeutic applications.

Small molecules binding to G-quadruplexes have been largely based on polycyclic planar aromatic compounds with at least one substituent terminating in a cationic group [16-18]. The original rationale for the planar moiety was that this would stack effectively onto planar G-quartets, which has been subsequently visualised in a number of crystallographic and

NMR studies of G-quadruplex-ligand complexes [19-24]. Suggestions that the ligands would intercalate between G-quartets that are in the interior of a quadruplex have not been supported by structural studies, which are unanimous in showing that such planar ligands stack onto a terminal G-quartet. The substituents are typically short acyclic chains (for example $-(CH_2)_3-$) terminating in a cationic nitrogen-containing group such as diethylamine, pyrrolidine or piperidine. The cationic charge requirement has led to the dogma that these groups reside in quadruplex grooves and directly contact phosphate groups. However the crystallographic evidence to date [23,24] indicates that these electrostatic contacts are rarely direct but are often mediated by bridging water molecules, with the space in the grooves containing structured water networks, analogous to those around duplex DNA sequences. The role of the loops in ligand recognition is less well understood. The TTA loops in a complex between a telomeric quadruplex and the trisubstituted acridine experimental drug BRACO-19 [23](Figure 1) have undergone a number of conformational changes compared to the TTA loops in the native quadruplex, so that they form pockets around the acridine substituents, notably around an aniline group. Thus the flexibility of loops containing at least three nucleotides may be a determining factor in the selectivity of ligand binding, although as yet there is insufficient structural data available for any rational design principles to be deduced.

Human telomeric quadruplex nucleic acids

Human telomeres comprise tandem repeats of the short DNA motif (TTAGGG) [25] together with an array of telomeric proteins [26,27•] as well as other more transient DNA-repair and damage-response proteins such as Ku [28]. Telomere function is to protect chromosomal ends from unwanted recombination and degradation. The terminal 150-250 nucleotides at the extreme 3' ends of telomeres are single-stranded [29], but protected by binding to multiple repeats of a single-stranded DNA binding protein, (hPOT1 in humans), which in turn interacts with other proteins in the core telomere complex, notably TPP1, to regulate telomerase action in cancer cells, and thus to maintain telomere length [30,31,32•]. Loss of hPOT1 deprotects telomeres, and initiates DNA damage-response mediated cell death. Small molecules that stabilise the single strand into a higher order (G-quadruplex) structure compete with hPOT1 and thus also initiate this response [33-35]. Thus quadruplex formation may itself be a DNA damage signal, producing analogous responses to other mediators of telomere damage [36,37,38•,39•] as do quadruplex-forming aptamers that also induce this response [40]. The biological function of induced telomeric quadruplexes remains to be fully clarified; an end-protective role has been suggested, there is evidence of functional interactions involving poly (ADP-ribose) polymerase-1 [41••], and in ciliates at least, quadruplex structure is involved in telomerase recruitment [42•]. However, to date there is no direct evidence of a role for telomeric G-quadruplexes in the functioning of telomeres in normal human cells.

Interest in telomeres and telomerase as therapeutic targets arose following the seminal finding that telomerase is expressed in ca 80-85% of cancer cells and primary tumours, but not in normal somatic cells [43]. Its function in the former is in large part to maintain telomere-length homeostatis (acting as a tumour promoter), whereas telomere shortening in the latter is in effect a tumour suppressor mechanism [27•]. Telomeres in somatic cells are gradually shortened as a consequence of the end-replication effect, and on reaching critically short lengths, cells then enter p53 and Rb-dependent replicative senescence, and ultimately apoptosis. The catalytic subunit of telomerase having reverse transcriptase enzymatic activity (hTERT in humans), synthesises TTAGGG repeats on to the end of the 3' single-stranded overhang. Inhibition of hTERT by siRNA, antisense or small-molecule inhibitors selectively inhibits cancer cell growth, and has been taken as a proof-of-principle that induction of telomere shortening is a viable therapeutic strategy [17].

The substrate for telomerase is the single-stranded G-rich telomeric DNA overhang, which is recognised by the RNA template component of the telomerase complex. Folding the substrate into a quadruplex structure was shown some years ago to inhibit catalytic activity [44]. The concept that this could be induced by small-molecule quadruplex stabilisation was first demonstrated using a disubstituted anthraquinone derivative [18]. A large number of quadruplex-binding ligands have been subsequently reported [16,17,45], although rather fewer have been evaluated in cell-based assays. The majority of G-quadruplex ligands in the literature contain a polycyclic heteroaromatic core, it is clear that this is not an essential requirement for quadruplex binding, since an increasing number of compounds that do not have this feature, of which the cyclic polyamine telomestatin (Figure 1) was the first such compound [46], nonetheless show both quadruplex affinity and telomerase inhibitory potency. More recent reports have demonstrated that acyclic non-conjugated compounds that are synthetically more accessible than telomestatin can have analogous levels of potency and selectivity (see for example Refs. [39,47-49]).

In vivo anticancer activity has been reported for a few telomeric quadruplex ligands (Figure 1), notably the trisubstituted acridine compound BRACO-19 [50], the polycyclic compound RHSP4 [38,51] and telomestatin [52]. For the future development of telomeric quadruplex ligands, it is noteworthy that it has been possible to demonstrate for BRACO-19 and RHSP4 that the *in vivo* target of these compounds is the telomeric DNA single-strand overhang, by the observations of hPOT1 and hTERT uncapping. None of these molecules has progressed beyond the experimental stage into clinical trial, possibly since these 1st-generation compounds are as yet insufficiently drug-like.

Telomere quadruplex-binding small molecules - mechanisms of action

The classic model of telomerase inhibition and consequent telomere attrition leading to senescence and apoptosis requires that cells with a mean telomere length of 5 kb, a 24-hour cell doubling time and a subsequent loss of ca 100 nucleotides per round of replication would reach critical telomere shortening in ca 40-50 days [53,54]. This was indeed the observation in dominant negative telomerase experiments, but would be therapeutically futile for cancer treatment. Initial findings using G-quadruplex ligands showed very different behavior, with senescence occurring within 7-10 days after cells were first treated and little evidence of concomitant telomere shortening. This has been subsequently shown to be characteristic of the class as a whole, and the observations of on-target *in vivo* activity are encouraging signs that clinical utility may be achievable with appropriate compounds.

Several quadruplex-binding ligands, such as the acridines BRACO-19 and RHSP4 induce replicative senescence in cancer cells with a few days of exposure (the effects on some cell lines are more immediate and may be telomere-length dependent). This activates the same DNA damage response that follows DNA double-strand breaks, and involves in particular ATM, p16^{INK4a} kinase and p53 pathways [37,38] which can be visualised by, for example, the appearance of characteristic DNA damage foci using an antibody to the damage-response protein γ H2AX [39], or by a significant population of cells undergoing end-to-end fusions in metaphase [55]. Such changes are analogous to those produced when the telomeric protein TRF2 is knocked out. This response is a consequence of the displacement of bound proteins from the single-stranded overhang, chiefly hPOT1 (Figure 2), as well as possible uncapping of telomerase from the ends. There are likely to be multiple mechanisms involved, some of which at least have crosstalk between them. For example, hPOT1 interacts with the telomeric protein Tpp1 and so facilitates telomere-length regulation by telomerase, and hPOT1 displacement dis-regulates telomerase function [31,32]. In addition, although the classic telomerase inhibition model does not appear to be followed by G-quadruplex-binding agents, most if not all cancer cells have marked telomere-length

heterogeneity, with some having extremely short (<1 kb) telomeres. It has been suggested that these cells are not only sensitive to senescence, as one would expect, but also their viability is critical to the cell population overall [56,57]. A Q-FISH technique has been used to determine that telomestatin is localised at telomeres during replication and has found that telomere replication is unaffected in mouse embryonic fibroblast (i.e. untransformed) cell lines [58]. The analogous experiments in cancer cells have not as yet been reported.

G-quadruplexes in gene promoters

There is now a growing body of work that has explored a hypothesis that links the existence of G-quadruplex-forming sequences in gene promoters to the transcriptional activity of the proximal gene. The 'promoter-quadruplex' hypothesis is shown in Figure 3 in its most general form and depicts the formation of G-quadruplex structures in gene promoters in a process that is temporally and mechanistically coupled to transcriptional activity. This hypothesis has inspired the suggestion that small molecules that interact with such DNA G-quadruplexes may exert functional changes by specifically modulating the expression of the gene(s) regulated by these promoter sequences. The prospect of small-molecule gene regulators that act via a G-quadruplex mechanism presents attractive opportunities for the design of novel therapeutic agents that target disease-related genes both in humans and in a wide range of anti-infective targets.

The proposed association between G-quadruplexes and transcription has its origins in studies that set out to explore the relationship between nuclease hypersensitivity sites, indicative of non-duplex DNA structure and the transcriptional activation of genes. A number of such nonduplex structures were found upstream of transcription start sites and close analysis of the sequence content of such regions revealed a prevalence of G-rich sequence. Indeed, many such sequence stretches would today be recognised as intramolecular G-quadruplex-forming motifs [59,60], and it was probably the chicken- β -globin gene that was the first to have a G-quadruplex sequence formally identified within its promoter [61].

A number of genome-wide computational studies have since been carried out to explore whether G-quadruplex sequence motifs are particularly prevalent near transcription start sites. A study that focused on human protein coding genes revealed that more than 40% of the genes had at least one G-quadruplex motif within 1 kb upstream of the transcription start site [62]. Furthermore, the quadruplex motifs within these upstream regions showed a strong positional bias towards the site at which transcription starts [62,63]. These features are analogous to those exhibited for 'classical' transcription factor binding regulatory sites [64] and would be consistent with G-quadruplex sequence motifs being *cis*-acting regulatory elements of transcription. Indeed, other computational studies have concluded that promoter-associated G-quadruplex motifs are proximal to [65] or overlapping with [66] transcription factor binding sites. Thus, suggesting a possible mechanism whereby G-quadruplex motifs may interact with transcription factor binding events. Although, such data do not distinguish whether it is G-quadruplex structure formation rather than the sequence *per se*, which is relevant to transcription.

Detailed investigations have been carried out on G-quadruplex motifs found within the promoters of a number of medically related genes, particularly proto-oncogenes. The first-studied and best-studied case, of relevance to this hypothesis, was that of the human *c-myc* for which a G-quadruplex-forming sequence was identified in a nuclease hypersensitive element upstream of the P1 promoter [67]. NMR spectroscopic studies have elucidated the complexity of G-quadruplex structures formed by the motif, with a well-defined G-quadruplex structure being obtained for a truncated sequence comprising just four sets of G-

tracts [68,69]. Genetic point mutations that destabilised the *c-myc* G-quadruplex structure, led to an increase in transcriptional activity, whereas the G-quadruplex interactive small-molecule TMPyP4 was found to reduce the transcriptional activity, supportive of the promoter-G-quadruplex hypothesis [70••].

Two distinct G-quadruplex-forming sequence motifs were identified in the core promoter of the human *c-kit* oncogenes [71,72]. Interestingly, both motifs flanked the binding site of the activating transcription factor sp1 by ~10 nucleotides. The structure formed by one of the *c-kit* G-quadruplex-forming motifs (*c-kit1*) was solved by NMR [7]. Single molecule Förster resonance energy transfer spectroscopy has suggested that the sequence of the second (*c-kit2*) G-quadruplex motif is sufficient to programme the formation of non-duplex structures within the context of an extended duplex [73]. An isoalloxazine small-molecule G-quadruplex ligand that binds both *c-kit* G-quadruplexes was shown to reduce the levels of *c-kit* mRNA in *c-kit* expressing cell line [74]. Studies have also confirmed the presence of a G-quadruplex-forming sequence within the promoter of *k-ras* [75], which is also sensitive to a reduction in transcriptional activity induced by the G-quadruplex interactive ligand TMPyP4. Experiments using nuclear extracts identified Ku, PARP-1 and hnRNP-A1 as proteins that bind the G-quadruplex-forming sequence [76••]. Interestingly, PARP-1 has also been found to be enzymatically activated upon binding to G-quadruplex DNA [41••], and hnRNP-A1 has been found to unwind G-quadruplex DNA [77], which may both be related to promoter-G-quadruplex function.

Promoter-G-quadruplexes have now been reported for a number of other cancer-related genes for which there is *in vitro* biophysical data, such as *VEGF*, *PDGF*, *HIF1a*, *bcl-2* and *RET* [78] and in some cases also preliminary chemical biology evidence in support of the 'promoter-G-quadruplex' hypothesis.

Conclusions and perspectives

Herein, we have focused our attention to DNA G-quadruplexes proposed to exist at telomeres and in the promoters of genes. There are indeed several other regions of the genome where the presence of G-quadruplex motifs have been noted, which include intragenic regions (both introns and exons) and the immunoglobulin class switch regions, where they are proposed to influence genetic rearrangements that ultimately alter the class of expressed antibody. It is also worth mentioning that stable G-quadruplex-forming sequences have been observed and studied in RNA transcripts, with respect to their role in untranslated regions associated with translation [79••] and alternative splicing events [80]. Whilst there is considerable opportunity (and need) to further explore the details of the structure, function and mechanisms associated with G-quadruplex motifs in all such genomic environments, the weight of current understanding has been built upon with respect to their incidence at telomeres and promoters.

The most compelling evidence for G-quadruplex function has been found at ciliate telomeres where telomere end-binding proteins have been shown to modulate G-quadruplex structure in a manner that is directly depending on a cell cycle-dependent phosphorylation of one of the proteins [81]. It has yet to be shown whether G-quadruplex structures are naturally formed and controlled in a similar manner at the telomeres of human cells. Past observations suggest that telomere-directed G-quadruplex ligands may capture and stabilise the G-quadruplex folded form, or alternatively induce G-quadruplex formation at the telomere, in either case leading to telomere uncapping and derailment of telomere maintenance with consequent antiproliferative effects such as the induction of a DNA damage response in affected cells. If indeed G-quadruplex formation and/or the accessibility of the telomere to small-molecule ligands are/is related to the cell cycle, this may render the

telomere of dividing cells differentially susceptible to small molecules and contribute to selectivity. A major challenge is to distinguish sequence, structural and biophysical properties of the human intramolecular telomeric G-quadruplex from the other G-quadruplexes at the molecular level, in principle facilitating the design of ligands selective for the human telomere, although there is as yet structural information on only a handful of quadruplexes (telomeric, and the *c-kit*, *c-myc* and *bcl-2* promoter quadruplexes), and general rules for predicting structure from sequence are in their infancy. There are some indications that the uniqueness of a quadruplex sequence within a genome can correspond to a unique three-dimensional structure, as in the instance of the *c-kitl* promoter quadruplex [82]. Also, several telomeric quadruplexes may be formed along a stretch of singlestranded telomeric DNA overhang, resulting in a quadruplex multimer as a target.

The evidence for the involvement of G-quadruplex sequence motifs in transcription has increased considerably during the past 10 years. As a consequence, the prospect of targeting such motifs to alter the transcriptional state of certain genes has become attractive [83]. The field stands to benefit from further, direct evidence of G-quadruplex *structure* formation in gene promoters in the context of a cellular, chromatinised genome. If indeed G-quadruplex structures form in promoter DNA, it is plausible that they have a transient existence, in which case it would be insightful to establish exactly when they form in relation to transcription events. There is also much to elucidate concerning the precise mechanistic relationship between promoter-G-quadruplex formation and the transcription machinery. Functional data from chemical biological studies [70•,74,75] support the hypothesis that G-quadruplex ligands are able to alter the expression of genes downstream of promoter-G-quadruplexes. However, it remains to be rigorously explored whether such ligands kinetically trap G-quadruplex DNA that pre-exists, or whether the ligands are capable of inducing G-quadruplex formation in cellular duplex DNA. Whilst there has been very good progress in this particular sub-domain of the G-quadruplex field, it remains a fertile area for further exploration. Continued chemical biological studies will provide fundamental insights into the nature of promoter-G-quadruplex motifs and the progress towards therapeutics will follow.

Molecular studies on the structure and recognition of G-quadruplex DNA have provided initial insights into the challenges associated with selective targeting by small molecules. Chemists in the field have already provided ample examples of small-molecule families that have shown a high degree of preference for G-quadruplex structures in favour of the DNA duplex. For therapeutics to become a reality, a level of specificity will be required between the various DNA G-quadruplexes that might be simultaneously present in a genome. The realisation that intramolecular G-quadruplexes can have distinct recognition features as a consequence of the length, sequence and folded topology of their loops, is a clear opportunity to be exploited for differential recognition by small molecules. The structural data showing selective interactions between fragments of small molecules and the loops of G-quadruplexes provide a source of inspiration for the design of future generations of small molecules that depend less on general G-tetrad recognition for binding. It is also plausible that alterations in the biological mechanism(s) which are causative of a disease phenotype (e.g. transcriptional upregulation of an oncogene in a particular cancer) may also have a characteristic structural alteration in the chromatin and/or DNA structure in the target region, which could improve target selectivity. It is thus possible that G-quadruplex-binding molecules with relatively modest selectivity between various G-quadruplexes, may still have utility in cancer therapeutics, provided they have low toxicity to normal cells. The compound quarfloxin is the first therapeutic agent designed to target quadruplexes to enter clinical evaluation and is currently in early phase II trials for carcinoid or neuroendocrine tumours. Its mode of action was initially suggested to involve *c-myc* promoter quadruplex stabilisation, but it is now proposed (www.cyclepharma.com) that it disrupts interaction

between ribosomal DNA G-quadruplexes and the abundant protein nucleolin. We anticipate that further advancement in this field will be followed by the discovery of more small molecules with therapeutic potential.

References and recommended reading

Papers of particular interest, published within the period of the review, have been highlighted as:

- of special interest
- of outstanding interest

1. Druker BJ. Translation of the Philadelphia chromosome into therapy for CML. *Blood*. 2009; 112:4808–4817. [PubMed: 19064740]
2. Hambley TW. Is anticancer drug development heading in the right direction? *Cancer Res*. 2009; 69:1259–1262. [PubMed: 19208831]
3. Dose C, Farkas ME, Chenoweth DM, Dervan PB. Next generation hairpin polyamides with (*R*)-3,4-diaminobutyric acid turn unit. *J Am Chem Soc*. 2008; 130:6859–6866. [PubMed: 18459783]
4. Davis JT. G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew Chem Int Ed*. 2004; 43:668–698.
5. Burge SE, Parkinson GN, Hazel P, Todd AK, Neidle S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*. 2006; 34:5402–5415. [PubMed: 17012276]
6. Patel DJ, Phan AT, Kuryavyi V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res*. 2007; 35:7429–7455. [PubMed: 17913750]
7. Phan AT, Kuryavyi V, Burge S, Neidle S, Patel DJ. Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J Am Chem Soc*. 2007; 129:386–4392.
8. Parkinson GN, Lee MPH, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*. 2002; 417:876–880. [PubMed: 12050675]
9. Luu KN, Phan AT, Kuryavyi V, Lacroix L, Patel DJ. Structure of the human telomere in K⁺ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J Am Chem Soc*. 2006; 128:9963–9970. [PubMed: 16866556]
10. Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res*. 2006; 34:2723–2735. [PubMed: 16714449]
11. Phan AT, Kuryavyi V, Luu KN, Patel DJ. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K⁺ solution. *Nucleic Acids Res*. 2007; 35:6517–6525. [PubMed: 17895279]
12. Lim KW, Amrane S, Bouaziz S, Xu W, Mu Y, Patel DJ, Luu KN, Phan AT. Structure of the human telomere in K⁽⁺⁾ solution: a stable basket-type G-quadruplex with only two G-tetrad layers. *J Am Chem Soc*. 2009; 131:4301–4309. [PubMed: 19271707] This determination of the solution structure of a novel fold for human telomeric quadruplexes shows that they have even more topological complexity than has been assumed.
13. Xue Y, Kan ZY, Wang O, Yao Y, Liu J, Hao YH, Tan Z. Human telomeric DNA forms parallel-stranded intramolecular G-quadruplex in K⁺ solution under molecular crowding conditions. *J Am Chem Soc*. 2007; 129:11185–11191. [PubMed: 17705383]
14. Phan AT, Kuryavyi V, Gaw HY, Patel DJ. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nat Chem Biol*. 2005; 1:167–173. [PubMed: 16408022]
15. Dai J, Chen D, Jones RA, Hurley LH, Yang D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res*. 2006; 34:5133–5144. [PubMed: 16998187]
16. Monchaud D, Teulade-Fichou MP. A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem*. 2008; 6:627–636. [PubMed: 18264563]

17. de Cian A, Lacroix L, Douarre C, Temime-Smaali N, Trentesaux C, Riou J-F, Mergny J-L. Targeting telomeres and telomerase. *Biochimie*. 2008; 90:131–155. [PubMed: 17822826]
18. Sun D, Thompson B, Cathers BE, Salazar M, Kerwin SM, Trent JO, Jenkins TC, Neidle S, Hurley LH. Inhibition of human telomerase by a G-quadruplex-interactive compound. *J Med Chem*. 1997; 40:2113–2116. [PubMed: 9216827]
19. Fedoroff OY, Salazar M, Han H, Chemeris VV, Kerwin SM, Hurley LH. NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry*. 1998; 37:12367–12374. [PubMed: 9730808]
20. Haider SM, Parkinson GN, Neidle S. Structure of a G-quadruplex-ligand complex. *J Mol Biol*. 2003; 326:117–125. [PubMed: 12547195]
21. Gavathiotis E, Heald RA, Stevens MF, Searle MS. Drug recognition and stabilisation of the parallel-stranded DNA quadruplex d(TTAGGG)₄ containing the human telomeric repeat. *J Mol Biol*. 2003; 334:25–36. [PubMed: 14596797]
22. Parkinson GN, Ghosh R, Neidle S. Structural basis for binding of porphyrin to human telomeres. *Biochemistry*. 2007; 46:2390–2397. [PubMed: 17274602]
23. Campbell NH, Parkinson GN, Reszka AP, Neidle S. Structural basis of DNA quadruplex recognition by an acridine drug. *J Am Chem Soc*. 2008; 130:6722–6724. [PubMed: 18457389]
24. Parkinson GN, Cuenca F, Neidle S. Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes. *J Mol Biol*. 2008; 381:1145–1156. [PubMed: 18619463]
25. Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, Wu JRA. Highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A*. 1988; 85:6622–6626. [PubMed: 3413114]
26. Blackburn EH. Switching and signaling at the telomere. *Cell*. 2001; 106:661–673. [PubMed: 11572773]
27. Stewart SA, Weinberg RA. Telomeres: cancer to human aging. *Annu Rev Cell Dev Biol*. 2006; 22:531–557. [PubMed: 16824017] An authoritative review on telomere biology and its relationship to human disease.
28. d'Adda di Fagnana F, Teo SH, Jackson SP. Functional links between telomeres and proteins of the DNA-damage response. *Genes Dev*. 2004; 18:1781–1799. [PubMed: 15289453]
29. Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev*. 1997; 11:2801–2809. [PubMed: 9353250]
30. Zaug AJ, Podell ER, Cech TR. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc Natl Acad Sci U S A*. 2005; 102:10864–10869. [PubMed: 16043710]
31. Denchi EL, de Lange T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature*. 2007; 448:1068–1071. [PubMed: 17687332]
32. Miyoshi T, Kanoh J, Saito M, Ishikawa F. Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length. *Science*. 2008; 320:1341–1344. [PubMed: 18535244] This paper defines the role of the POT1 protein in telomere maintenance.
33. Gomez D, O'Donohue MF, Wenner T, Douarre C, Macadre J, Koebel P, Giraud-Panis MJ, Kaplan H, Kolkes A, Shin-Ya K, Riou J-F. The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences in vitro and induces GFP-POT1 dissociation from telomeres in human cells. *Cancer Res*. 2006; 66:6908–6912. [PubMed: 16849533]
34. Gomez D, Wenner T, Brassart B, Douarre C, O'Donohue MF, El KV, Shin-Ya K, Morjani H, Trentesaux C, Riou J-F. Telomestatin-induced telomere uncapping is modulated by POT1 through G-overhang extension in HT1080 human tumor cells. *J Biol Chem*. 2006; 281:38721–38729. [PubMed: 17050546]
35. Gunaratnam M, Greciano O, Martins C, Reszka AP, Schultes CM, Morjani H, Riou J-F, Neidle S. Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochem Pharmacol*. 2007; 74:679–689. [PubMed: 17631279]

36. d'Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, Von Zglinicki T, Saretzki G, Carter NP, Jackson SP. A DNA damage checkpoint response in telomere-initiated senescence. *Nature*. 2003; 426:194–198. [PubMed: 14608368]
37. Tauchi T, Shin-Ya K, Sashida G, Sumi M, Nakajima A, Shimamoto T, Ohyashiki JH, Ohyashiki K. Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene*. 2003; 22:5338–5347. [PubMed: 12917635]
38. Salvati E, Leonetti C, Rizzo A, Scarsella M, Mottolese M, Galati R, Sperduti I, Stevens MF, D'Incalci M, Blasco M, et al. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J Clin Invest*. 2007; 117:3236–3247. [PubMed: 17932567] A study of the mode of action of the quadruplex-binding compound RHPS4 and its effects following telomere binding.
39. Rodriguez R, Müller S, Yeoman JA, Trentesaux C, Riou J-F, Balasubramanian S. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J Am Chem Soc*. 2008; 130:15758–15759. [PubMed: 18975896] A demonstration that a G-quadruplex ligand produces effects on telomere integrity and DNA damage responses in human cells.
40. Qi H, Lin CP, Fu X, Wood LM, Liu AA, Tsai YC, Chen Y, Barbieri CM, Pilch DS, Liu LF. G-quadruplexes induce apoptosis in tumor cells. *Cancer Res*. 2006; 66:11808–11816. [PubMed: 17178877]
41. Soldatenkov VA, Vetcher AA, Duka T, Ladame S. First evidence of a functional interaction between DNA quadruplexes and poly(ADP-ribose) polymerase-1. *ACS Chem Biol*. 2008; 3:214–219. [PubMed: 18338862] This study demonstrates that this protein not only binds to quadruplexes *in vitro* with high affinity, but also is catalytically activated by this interaction, both by promoter and telomeric quadruplexes.
42. Paeschke K, Juranek S, Simonsson T, Hempel A, Rhodes D, Lipps HJ. Telomerase recruitment by the telomere end binding protein-beta facilitates G-quadruplex DNA unfolding in ciliates. *Nat Struct Mol Biol*. 2008; 15:598–604. [PubMed: 18488043] It has been previously established that telomeric G-overhangs of the ciliate *Stylonychia lemnae* fold into a G-quadruplex DNA structure *in vivo*. This study defines the role of phosphorylation of the telomere end-binding protein TEBPβ in the regulation of quadruplex unfolding.
43. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW. Specific association of human telomerase activity with immortal cells and cancer. *Science*. 1994; 266:2011–2015. [PubMed: 7605428]
44. Zahler AM, Williamson JR, Cech TR, Prescott DM. Inhibition of telomerase by G-quartet DNA structures. *Nature*. 1991; 350:718–720. [PubMed: 2023635]
45. Ou T, Lu Y, Tan J, Huang Z, Wong K, Gu L. G-quadruplexes: targets in anticancer drug design. *ChemMedChem*. 2008; 3:690–713. [PubMed: 18236491]
46. Kim MY, Vankayalapati H, Shin-Ya K, Wierzbicka K, Hurley LH. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J Am Chem Soc*. 2002; 124:2098–2099. [PubMed: 11878947]
47. Moorhouse AD, Santos AM, Gunaratnam M, Moore M, Neidle S, Moses JE. Stabilization of G-quadruplex DNA by highly selective ligands via click chemistry. *J Am Chem Soc*. 2006; 128:15972–15973. [PubMed: 17165715]
48. Drewe WC, Nanjunda R, Gunaratnam M, Beltran M, Parkinson GN, Reszka AP, Wilson WD, Neidle S. Rational design of substituted diarylureas: a scaffold for binding to G-quadruplex motifs. *J Med Chem*. 2008; 51:7751–7767. [PubMed: 19053833]
49. Dash J, Shirude PS, Balasubramanian S. G-quadruplex recognition by bis-indole carboxamides. *Chem Commun*. 2008:3055–3057.
50. Burger AM, Dai F, Schultes CM, Reszka AP, Moore MJ, Double JA, Neidle S. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Res*. 2005; 65:1489–1496. [PubMed: 15735037] Demonstration that the acridine compound BRACO-19 has antitumour activity in a xenograft model for a human cancer.
51. Phatak P, Cookson JC, Dai F, Smith V, Gartenhaus RB, Stevens MF, Burger AM. Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro

and in vivo consistent with a cancer stem cell targeting mechanism. *Br J Cancer*. 2007; 96:1223–1233. [PubMed: 17406367] Along with Ref [50•], this paper shows that the loss of hTERT occurs during *in vivo* treatment of a cancer xenograft with a G-quadruplex ligand, suggesting that this may be a suitable biomarker for future clinical trials of such agents.

52. Tauchi T, Shin-Ya K, Sashida G, Sumi M, Okabe S, Ohyashiki JH, Ohyashiki K. Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. *Oncogene*. 2006; 25:5719–5725. [PubMed: 16652154]
53. Kelland LR. Targeting the limitless replicative potential of cancer: the telomerase/telomere pathway. *Clin Cancer Res*. 2007; 13:4960–4963. [PubMed: 17785545]
54. Oganessian L, Bryan TM. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *Bioessays*. 2007; 29:155–165. [PubMed: 17226803]
55. Incles CM, Schultes CM, Kempinski H, Koehler H, Kelland LR, Neidle S. A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. *Mol Cancer Ther*. 2004; 3:1201–1206. [PubMed: 15486186]
56. Hemann MT, Strong MA, Hao LY, Greider CW. The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell*. 2001; 107:67–77. [PubMed: 11595186]
57. Londoño-Vallejo JA. Telomere instability and cancer. *Biochimie*. 2008; 90:73–82. [PubMed: 17728038]
- 58•. Arnoult N, Shin-Ya K, Londoño-Vallejo JA. Studying telomere replication by Q-CO-FISH: the effect of telomestatin, a potent G-quadruplex ligand. *Cytogenet Genome Res*. 2008; 122:229–236. [PubMed: 19188691] This study shows directly that telomestatin bind to telomeres, by affecting the efficiency of FISH telomere probes. It also finds that this compound does not affect telomere replication in normal cells.
59. Todd AK, Johnston M, Neidle S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res*. 2005; 33:2901–2907. [PubMed: 15914666]
60. Huppert JL, Balasubramanian S. Prevalence of quadruplex sequences in the human genome. *Nucleic Acids Res*. 2005; 33:2908–2916. [PubMed: 15914667]
61. Howell RM, Woodford KJ, Weitzmann MN, Usdin K. The chicken beta-globin gene promoter forms a novel “cinched” tetrahelical structure. *J Biol Chem*. 1996; 271:5208–5214. [PubMed: 8617804]
- 62•. Huppert JL, Balasubramanian S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res*. 2007; 35:406–413. [PubMed: 17169996] This computational study showed that G-quadruplex motifs were concentrated in the promoter (1 kb upstream) region of human protein coding genes, with 43% of genes having one or more such motif in the promoter. The paper also demonstrated a general positional bias of G-quadruplex motifs towards the transcription start sites.
- 63•. Verma A, Halder K, Halder R, Yadav VK, Rawal P, Thakur RK, Mohd F, Sharma A, Chowdhury S. Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved *cis*-regulatory elements in human and related species. *J Med Chem*. 2008; 51:5641–5649. [PubMed: 18767830] This paper presents computational data in support of the enrichment of G-quadruplex motifs immediately upstream and downstream of the transcription start sites of genes in the genomes of human, chimpanzee, rat and mouse. Those motifs also showed a significant level of conservation.
64. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature*. 2005; 434:338–345. [PubMed: 15735639]
65. Todd AK, Neidle S. The relationship of potential G-quadruplex sequences in cis-upstream regions of the human genome to SP1-binding elements. *Nucleic Acids Res*. 2008; 36:2700–2704. [PubMed: 18353860]
- 66•. Eddy J, Maizels N. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res*. 2008; 36:1321–1333. [PubMed: 18187510] This study showed that G-quadruplex motifs are enriched in the first introns (i.e. downstream) of human genes. The paper also concludes that G-quadruplex motifs upstream of transcription start sites overlap with predicted transcription factor recognition sites.

67. Simonsson T, Pecinka P, Kubista M. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* 1998; 26:1167–1172. [PubMed: 9469822]
68. Phan AT, Modi YS, Patel DJ. Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J Am Chem Soc.* 2004; 126:8710–8716. [PubMed: 15250723]
69. Ambrus A, Chen D, Dai J, Jones RA, Yang D. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry.* 2005; 44:2048–2058. [PubMed: 15697230]
- 70••. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A.* 2002; 99:11593–11598. [PubMed: 12195017] The first example of a small-molecule G-quadruplex ligand that alters the transcriptional activity of a gene (c-myc) with a promoter-G-quadruplex motif as a target. The study also demonstrated that G-quadruplex-destabilising mutations in the promoter also lead to altered transcription. Both types of experiments were consistent with a mechanism whereby the c-myc G-quadruplex was acting as a repressor element.
71. Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK, Ladame S, Balasubramanian S, Neidle S. Putative DNA quadruplex formation within the human c-kit oncogene. *J Am Chem Soc.* 2005; 127:10584–10589. [PubMed: 16045346]
72. Fernando H, Reszka AP, Huppert J, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry.* 2006; 45:7854–7860. [PubMed: 16784237]
73. Shirude PS, Okumus B, Ying L, Ha T, Balasubramanian S. Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene c-kit. *J Am Chem Soc.* 2007; 129:7484–7485. [PubMed: 17523641]
74. Bejugam M, Sewitz S, Shirude PS, Rodriguez R, Shahid R, Balasubramanian S. Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression. *J Am Chem Soc.* 2007; 129:12926–12927. [PubMed: 17918848]
75. Cogoi S, Xodo LE. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.* 2006; 34:2536–2549. [PubMed: 16687659]
- 76••. Cogoi S, Paramasivam M, Spolaore B, Xodo LE. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res.* 2008; 36:3765–3780. [PubMed: 18490377] This paper describes the isolation of three nuclear proteins (hnRNP-A1, Ku70 and PARP-1) that each recognise the G-quadruplex folded form of the k-ras. It is suggested that these proteins, along with the k-ras promoter-G-quadruplex element, may be involved in the regulation of transcription.
77. Paramasivam M, Membrino A, Cogoi S, Fukuda H, Nakagama H, Xodo LE. Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription. *Nucleic Acids Res.* 2009 doi: 10.1093/nar/gkp138. [Epub ahead of print].
78. Qin Y, Hurley LH. Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie.* 2008; 90:1149–1171. [PubMed: 18355457]
- 79••. Kumari S, Bugaut A, Huppert JL, Balasubramanian S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat Chem Biol.* 2007; 3:218–221. [PubMed: 17322877] This paper describes an RNA G-quadruplex that forms proximal to the 5' cap of the UTR of NRAS. Functional studies showed that the RNA G-quadruplex inhibited the translation of the downstream coding region.
80. Gomez D, Lemarteleur T, Lacroix L, Mailliet P, Mergny J-L, Riou J-F. Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res.* 2004; 32:371–379. [PubMed: 14729921]
81. Paeschke K, Simonsson T, Postberg J, Rhodes D, Lipps HJ. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat Struct Mol Biol.* 2005; 12:847–854. [PubMed: 16142245]
82. Todd AK, Haider SM, Parkinson GN, Neidle S. Sequence occurrence and structural uniqueness of a G-quadruplex in the human c-kit promoter. *Nucleic Acids Res.* 2007; 35:5799–5808. [PubMed: 17720713]

83. Sun D, Hurley LH. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J Med Chem.* 2009 doi: 10.1021/jm900055s.

Figure 1.
Structures of three quadruplex-forming ligands for which *in vivo* data have been reported.

figure 2.

Schematic of the cellular consequences of inducing quadruplex formation by means of a small-molecule ligand within the single-stranded telomeric DNA overhang at the 3' end of a cancer cell chromosome. The single-strand binding protein hPOT1 is displaced from the overhang by quadruplex formation, so that recognition of the overhang by the RNA subunit of telomerase (hTR) action is impaired, telomerase-catalysed synthesis of telomeric DNA repeats is inhibited and telomeric DNA is shortened. The unmasking of the overhang from associated proteins can also lead to a DNA damage response.

Figure 3.

Schematic of the promoter-G-quadruplex hypothesis. G-quadruplex-forming sequence motifs in the upstream (promoter) region (in green) of genes (in red) may fold into G-quadruplex structures. The formation of G-quadruplex, rather than duplex, DNA structure in the promoter is associated with an altered state of transcription. This hypothesis would suggest that any molecule capable of interacting with the specific G-quadruplex could modulate the transcriptional activity of the associated gene.